OPIC
Office de la propriété intellectuelle du Canada

CIPO
Canadian Intellectual Property Office

(12) (19) (CA) **Demande-Application**

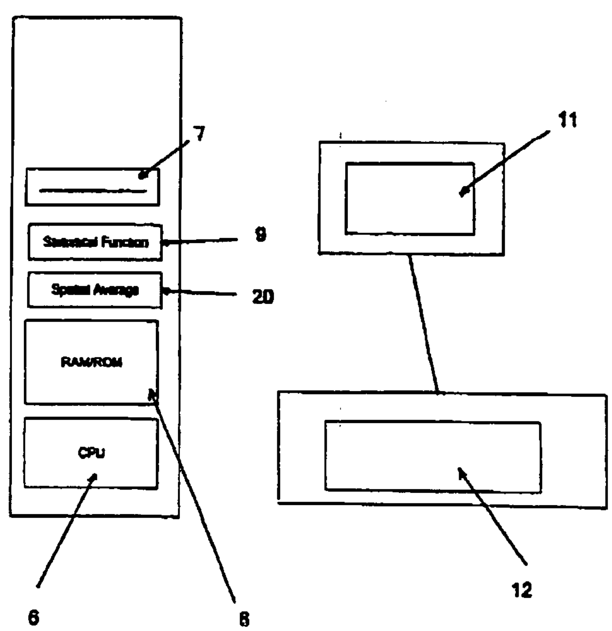
(21) (A1) **2,293,174**
(22) 1999/12/30
(43) 2000/06/30

(72) WACHTER, SUSAN M., US
(72) GILLEN, KEVIN, US
(72) AMOS, PAUL, US
(71) THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA, US

(51) Int.Cl.^{7} G06F 17/60, G11B 23/00
(30) 1998/12/30 (60/114,248) US
(30) 1999/12/29 (UNKNOWN) US

(54) **METHODE ET APPAREIL PERMETTANT DE PREDIRE LE PRIX DE TRANSACTION D'UNE PROPRIETE**

(54) **METHOD AND APPARATUS FOR PROPERTY TRANSACTION PRICE PREDICTION**



(57) The present invention provides a method and computer-executable instructions on a computer-readable medium for predicting a property transaction price for a prospective transaction. The method determines a fast value and a second value and predicts the property transaction as a function of those values. The first value is based on one or more attributes of the prospective transaction location and the second value is based on a relation of the prospective transaction location to one or more other properties.

METHOD AND APPARATUS FOR PROPERTY TRANSACTION PRICE PREDICTION

This application claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Patent Application Serial No. 60/114,248, which was filed on December 30, 1998 and is hereby incorporated by reference in its entirety.

Field of the Invention

This invention relates to a method and apparatus for automated prediction of real estate transaction prices using statistical models.

Background of the Invention

Fluctuations in the price of goods and services are typical in free market economic systems. Fluctuating market valuations imply that market participants bear certain risks. As a result, there is a constant need for improved risk management techniques and risk reduction methods. Banks and other mortgage lenders are a familiar example of market participants that use such techniques on a daily basis. Also, mortgage lenders must endeavor to make statistically accurate risk assessments when evaluating potential loan applicants. This assessment must necessarily include an estimation of the fair value of the property to be mortgaged. Thus, in an effort to determine the expected price of a property transaction, mortgage lenders, real estate brokers and others in need of property sale prediction conventionally have used automated statistical prediction models.

Traditionally, automated statistical prediction models have been based on two types of data: hedonic characteristics and neighboring property sales. A hedonic characteristic is a certain feature of a property whose price is being predicted. For example, a property's square footage, style of architecture, amenities, number of bedrooms, number of stories, type of exterior siding, or existence of a swimming pool are physical hedonic features. In addition, characteristics of a property's surrounding neighborhood, including its school district's average Scholastic Aptitude Test (SAT) scores, median income, or crime rates are attributes that affect a property's value. The following formula has been used to model the linear statistical relationship of a property's value as a function of its characteristics and attributes (*i.e.*, its hedonic characteristics):

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j H_{ji} + \varepsilon_i \qquad \varepsilon_i \sim iid(0,\sigma^2)\forall\, i = 1,2,\ldots,N$$

Where:

β_0 is approximately zero;

β_j = the jth hedonic characteristic;

N = the total number of properties in the dataset;

y_i = transaction price of the ith property;

H_{ji} = the value of the jth hedonic characteristic for the ith property; and

ε_i = the residual for each observation.

In this equation, $\varepsilon_i \sim iid(0,\sigma^2)\forall\, i = 1,2,\ldots,N$ means that the ε's are distributed independently and identically (*i.e.*, "iid") with a mean of zero and a finite variance for all N observations. However, the second type of data commonly used in automated statistical

prediction models, namely neighboring property sales, violates the assumption that the ε_i's are distributed independently, and thus has a negative impact on the accuracy of the hedonic estimations.

Neighboring property sales data is used to predict the price of a property based on prior property sales in the same jurisdiction. Because neighborhoods are typically characterized by local homogeneity, the probability that a neighboring home is very similar both in physical characteristics and value is very high. The implication of local homogenity is that the value of a given property is not completely independent of the values of surrounding properties. Consequently, the measurement error, or ε_n, associated with the model's predicted price for a given property exhibits spatial dependence. Thus, the assumption that the ε_i's are distributed independently is violated, with negative consequences for the correct estimation of the β's. A further implication of local homogeneity is that the influence of a given property declines with distance.

Currently, however, systems directed to prediction of property transaction prices are not optimally accurate. Empirical observation has shown that neighboring property sales within a geographic vicinity are perhaps the most accurate predictors of property transaction price. However, conventional prediction models based on neighboring property sales within a jurisdiction may not fully capture the predictive capability of observations within a geographic vicinity range. For example, relevant property sales in close proximity may fall outside defined jurisdictional borders and thus be excluded from the prediction system. Also, conventional prediction models have failed to combine detailed hedonic characteristic models with neighboring property sales models. Finally, to the extent current price prediction systems exhibit a degree of accuracy, such accuracy

comes as a result of basing prediction on numerous statistical variables, making the maintaining of neighboring property sale data a time consuming and costly task.

Therefore, it would be advantageous to provide a property transaction system that fully captures the predictive capability of neighboring property sales over a distance at which properties exhibit significant spatial covariance, while incorporating significant hedonic characteristics of the subject property. In addition, it would be advantageous to provide a property transaction system that achieves maximally accurate predictions using a minimum of observed statistical variables.

Summary of the Invention

The present invention provides a method and computer-executable instructions on a computer-readable medium for predicting a property transaction price for a prospective transaction. The method determines a first value and a second value and predicts the property transaction as a function of those values. The first value is based on one or more attributes of the prospective transaction location and the second value is based on a relation of the prospective transaction location to one or more other properties. Specifically, the first value is determined by selecting one or more attributes of the prospective transaction location, determining an attribute value for each of the attributes of the prospective transaction location, and adding the attribute values. The second value is determined by first determining one or more categorical ranges relative to the prospective transaction location, for example discrete, sequential and relevant ranges of distance from the prospective transaction location. Average property values within each categorical range are then calculated using a spatial average algorithm. Each of the average property

values are weighted based on their associated categorical range, and the weighted average property values are added.

Brief Description of Drawings

The features and advantages of the invention will be apparent from the following detailed description in conjunction with the attached drawings, of which:

Figure 1 is a block diagram representing a computer system in which aspects of the present invention may be incorporated;

Figure 2 is a flowchart showing a process for predicting a transaction price of a prospective property transaction, in accordance with one embodiment of the present invention;

Figures 3A and 3B depict a flowchart showing in greater detail the spatial algorithm process depicted in Figure 2;

Figure 4 depicts an input and output data set of the spatial algorithm process depicted in Figures 3A and 3B for a three observations input data set;

Figure 5 shows a sample computer generated graphical map with observational data represented as points, and distance ranges relative to a prospective transaction location represented as concentric circles; and

Figure 6 is a zoomed-in view of the map shown in Figure 5.

Detailed Description of the Preferred Embodiment

The present invention overcomes the above limitations in the prior art by providing a method for predicting a property transaction price for a prospective transaction location

as a function of a first value based on attributes of the subject property, and a second value based on a relation to neighboring property sales. In particular, the second value is determined by a spatial algorithm that computes the average price of the neighboring properties for different categorical distances.

Figure 1 shows an overview of a computer system environment in which the present invention may be wholly or partially employed. A computer 1 includes a computer processing device 6, storage device 7, memory device 8, display device 11 and user input device 12. Computer processing device 6 can be implemented with, for example, a single microprocessor chip, printed circuit board, several boards or other devices. Storage device 7 can be implemented with, for example, internal hard disks, tape cartridges, or CD-ROMs. Memory device 8 can be implemented with, for example, a collection of random access memory (RAM) and read only memory (ROM) chips. Display device 11 can be implemented with any display, for example, a monitor. User input device 12 can be implemented with, for example, a keyboard, mouse or scanner.

In accordance with the presently described embodiment, computer 1 also includes a statistical function element 9 and a spatial computation element 20, both of which comprise instructions executed by computer processing device 6. Statistical function element 9 includes one or more processes of the sort typically included in statistical manipulation software (*i.e.*, performing re-formatting, semivariogram estimation and stepwise regression operations on a given set of input data). Given a prospective transaction, and prior transaction data in the same vicinity, spatial computation element 20 calculates the average transaction price (over a user specified period) for real estate within each of a plurality of vicinity ranges calculated from the location of the prospective

transaction. Spatial computation element 20 instructions may be written in any conventional programming language, such as Arcview's Avenue™, or other types of conventional Geographic Information System (GIS) software.

Figure 2 is a flowchart showing a process for predicting a transaction price of a prospective property, in accordance with one embodiment of the present invention. As shown in Figure 2, beginning with a transactions dataset of *N* property values that contains geographic coordinates and sales price of the properties, the method of the invention models the spatial structure of the first- and second-order moments of local property values in accordance with the following six steps. It should be appreciated that the distance of the properties to the subject property is just one example of a relevant relation. Others relevant relations may include a comparison of the type of property, size of the property, and other amenities found on the respective properties (*e.g.*, swimming pool and type of exterior siding).

In step 100, a property level transaction dataset is obtained. The sole input to this process is the property-level transactions dataset that contains geographic coordinates and sales price of each observation. For example, the geographic coordinates may be the latitude and longitude of each observation, and the observations may be single-family homes. Several rules govern the selection of observations to include in the input dataset:

1) Homogeneity of property type - This process assumes that a separate model is being estimated for each type, or class, of properties. Consequently, it would be incorrect to include multifamily properties or condominiums in a dataset of single-family homes.

2) Short time-span - The data should be for a finite, and presumably recent, period of time. If the goal of estimating a spatial model of property valuation is to predict the sales price of properties that are expected to transact in the *near* future, then each observation in the data should be a transaction that has occurred no more than a few years ago.

3) Finite degree of localness - The entire dataset should be properties from a fixed and known locality. It would be incorrect to use observations from a very large region, such as an entire U.S. State. Nothing larger than a county and nothing smaller than a neighborhood is the suggested feasible range.

An example of a dataset that would provide the conditions for an ideal and robust estimation would be one that contained the population of all single-family house sales for a particular city from the last year. Large and recent random samples from a particular county or municipality are also viable. Both sales price and spatial coordinates should be measured as accurately and completely as possible.

In step 110, the data should be is read into a personal computer (PC) from its raw text file format and formatted in the PC. A PC with a processor speed of at least 200 MHz and a RAM capacity of 128 Mb is recommended. If the file is large, such as 500,000 KB for a large U.S. city, then sufficient disk space capacity should be installed. Traditional data-manipulation software applications such as SAS™, Matlab™, or Excel™ are recommended. First, any necessary data formatting and cleaning to correct flawed observations and variables must be performed. For example, latitude and longitude should

be in decimal degrees, and sales price should be in dollars. Any observations with missing or inaccurate values that cannot be corrected should be eliminated. The cleaned dataset should be a robust representation of a particular type of recent property sales in a particular locality.

It is generally desirable to take transformations of the geographic coordinates to input into the model as independent variables. Such transformations are useful for capturing any first-order global trends in property values within a given locality. If no such first-order effects exist (*i.e.*, the data is covariance stationary), then they would not be statistically significant in the stepwise regressions estimated in step 150 of the process, and thus dropped. Some suggested transformations include:

$$X_i^n;$$
$$Y_i^n;$$
$$X_i^n Y_i^n;$$
$$Ln\left(X_i^n\right);$$
$$Ln\left(Y_i^n\right); and$$
$$Ln\left(X_i^n Y_i^n\right)$$

Where: X_i = X-coordinate of the *i*th observation (*e.g.*, longitude)

Y_i = Y-coordinate of the *i*th observation (*e.g.*, latitude)

$n = ½, 1, 2, 3, \ldots$

In step 120, the maximum distance *D* at which properties no longer exhibit any significant spatial covariance must be determined. While there are a variety of existing statistical methods to accomplish this, it is recommend that the explicit estimation of a semivariogram be performed by a software package such as SAS™ or Matlab™. The advantage of the semivariogram approach is that it explicitly computes the *D* = 'range' at which covariance between properties for a particular locality is no longer significant, or is close to zero. A description of explicit functional forms can be found in the literature. *See Interactive Spatial Data Analysis*, Trevor C. Bailey and Anthony C. Gatrell, Addison Wesley Longman Ltd., 1995; *Statistics for Spatial Data*, Noel A.C. Cressie, John Wiley & Sons, Inc., 1993.

Other more traditional methods include the class of autocovariance and autocorrelation estimators. Ultimately, a value of *D* must be decided upon to determine the maximum range at which average property values will be computed. Typically, *D* will be no more than a few miles.

In step 130, the researcher determines the number *j* and width of categorical ranges at which average property values will be computed. The categorical ranges may represent ranges of distance from the subject property, for example. Depending upon the processor and memory constraints on the user's computing system, the user may employ relatively small random samples of the data and the given spatial average algorithm to experiment

with the value of j and the spatial width of each category to determine which configuration is optimal. Then, for a given configuration, the researcher may regress sales price on the j independent variables and examine the sign, value and significance of the parameter. However, an ideal decision rule is that the sum of the j parameter coefficients of the regression be relatively close to 1.

More formally:

$$\sum_{k=1}^{j} \lambda_k = (1 \pm \delta)$$

$$y_i = \lambda_0 + \sum_{k=1}^{j} \lambda_k V_{ki} + \varepsilon_i \qquad \varepsilon_i \sim iid(0,\sigma^2) \forall i = 1,2,\ldots, M$$

Where:

M = the number of properties in the random sample dataset ($M<N$);

y_i = transaction price of the ith property;

λ_k = regression coefficient for the kth categorical distance;

V_{ki} = the average value of all properties within some distance < D mile(s), for the ith property;

ε_i = the residual for ith observation;

δ = an arbitrarily small value, relative to 1 (*e.g.*, .005); and

λ_0 should not be significantly different from zero.

As an example, consider the following results where D=1.5 miles, j=4 and δ=.0002 was found to be a successful parameterization:

$$y_i = 0.0021 + .87(V_{1i}) + .075(V_{2i}) + .0357(V_{3i}) + .0195(V_{4i})$$

$$\sum_{k=1}^{4} \lambda_k = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = .87 + .075 + .0357 + .0195 = 1.0002$$

Where:

y_i = transaction price of the ith property;

λ_k = regression coefficient for the kth categorical distance;

V_{1i} = the average value of all properties within 1/4 mile, for the ith property;

V_{2i} = the average value of all properties beyond 1/4 mile but within ½ mile, for the ith property;

V_{3i} = the average value of all properties beyond ½ mile but within 1 mile, for the ith property;

V_{4i} = the average value of all properties beyond 1 mile but within 1.5 miles, for the ith property;

λ_0 is approximately equal to zero.

Intuitively, this estimation takes a given property value to be a weighted average of surrounding property values, where distance is the weight.

In step 140, having decided upon the values of D, j, and the width of each jth interval, the researcher then executes the pre-programmed spatial average algorithm to compute the average price of surrounding properties in each jth interval, for each ith subject property. The program then attaches the j fields to the original dataset, populated with the values computed by the algorithm. This program will be described in greater detail below with respect to Figures 3A and 3B.

In step 150, to decide upon a final, parsimonious specification, a stepwise

regression algorithm is used that estimates the model with different subsets of variables, adding or removing variables at each iteration for a given statistical significance criteria. Use of such an algorithm facilitates convergence to a particular subset f explanatory variables that have the highest predictive power and also comply with underlying statistical requirements.

In step 160, the final output is an equation that can be utilized to predict the sale price for out-of-sample properties which have not transacted. The equation will be of the form:

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j H_{ji} + \sum_{j=1}^{m} \lambda_j V_{ji} + \varepsilon_i \qquad \varepsilon_i \sim iid(0,\sigma^2)\,\forall\, i = 1,2,\ldots,N$$

Where:

N = the total number of properties in the dataset;

β_0 is approximately zero;

β_j = regression coefficient for the jth hedonic characteristic;

y_i = transaction price of the ith property;

H_{ji} = the value of the jth hedonic characteristic (*e.g.*, latitude, longitude, etc.) for the ith property;

V_{ji} = the average value of all properties within a specified categorical distance, for the ith property;

λ_j = regression coefficient for the jth categorical distance; and

ε_i = the residual for each observation;

The H_n variables are simply the values of the geographic coordinates (*e.g.*, latitude and longitude) and/or transformations of these values (as discussed with regard to step 110). Thus, after the β_j's and λ_j's have been estimated (as discussed above), the equation can be used to predict the price of an out-of-sample property (the *y*) that has not yet transacted by entering the values of the coordinate characteristics (the H_j's) and the values of the spatial averages (the V_j's). Notably, the geographic coordinates are included in the regression to capture *global* (i.e., *first-order*) effects in the data, and the spatial averages are included in the regression to capture *local* (i.e., *second-order*) effects. Stated differently, the values of β_j's represent measures of how the characteristics of a particular property affects its value, and the λ_j's represent measures of how the value of surrounding properties affects a particular property's value. Thus, the combination of hedonic values and neighboring property sales are fully captured in the predictive model, as expressed in the above equation.

A regression coefficient measures the change in the value of the dependent variable given a unit change in the independent variable tied to that coefficient. A regression equation measures the statistical relationship between variation in the dependent variables and variation in the independent variables. Notationally, a regression equation measures: $E[Y_i \mid X=X_i]$. This is the expected value of Y_i, given that X equals X_i. In this case, the value of Y_i (the price of property i) is being predicted, given that X (the Average Value within 1/8 mile of property i) takes on some particular value X_i.

In the above equation, regression coefficients, β and λ, may be estimated to explicitly measure how variation in latitude, longitude, and local prices affects variation in property values. For example, if "price" were simply regressed on "longitude" for

Philadelphia, a value f β=2300 may result. Because the value of longitude increases from west to east, this measures the change in average property values as the location of the property moves in 1-degree increments from West Philadelphia to Center City Philadelphia. Literally interpreted, β=2300 implies that average property values are increasing by $2,300 for every 1 degree increase in longitude. This result actually comports with common understanding of property values in Philadelphia. Additionally, if "price" is simply regressed on the "average value of all properties within 1/8 of a mile" from Philadelphia, the result might be a value of λ=.96. Literally interpreted, this implies that a $1 increase in the average price of surrounding properties within 1/8 of a mile causes a $0.96 cent increase in the price of the subject property. This should again seem intuitive, since increasing prices of surrounding properties imply that the neighborhood is becoming more desirable to live in, thus increasing the value of property within a neighborhood.

Figures 3A and 3B present a flowchart showing in greater detail the spatial algorithm process depicted in Figure 2. Beginning with a property-level dataset of property location and transactions price, this algorithm computes the average transaction prices of surrounding properties for a set of categorical distances, and adds these values as *j* new fields to the original dataset.

The algorithm begins in step 200 with an N-by-3 column dataset of property transactions. N is the number of property observations in the dataset. The first two columns of the data may be the spatial X-Y coordinates of each property, for example latitude and longitude in decimal degrees. The third column may be the transaction price of the property, for example in currency format. In one embodiment, the data file is in a tab-delimited text (*e.g.*, *.txt) format, and column headers are the first row of the dataset.

It is preferable, but not necessary, that the data be sorted on the X-Y coordinate fields.

In step 210, a file is read and a loop is commenced for k=1 though N observations. The file may be read, for example, using the Arcview™ application environment. The data is added to a view window as a spatial theme by choosing "View->Add Event Theme" from the menu bar. Arcview™ will then let the user choose from all the data tables available in this project file, so the user need only double-click on the table name. If column headers are the first row of the dataset, Arcview™ will automatically choose latitude and longitude as the X-Y coordinates of each property. The data will now appear as a point theme on the map of the particular locality in which the properties are located. It is assumed that the user has already defined the maximum distance D at which property values are correlated, and the width and number of the j categorical distance rings. For example:

D=2 miles;

j=5 categorical distances;

1st categorical distance = 0 to 1/8 miles;

2nd categorical distance = 1/8 to 1/4 miles;

3rd categorical distance = 1/4 to 1/2 miles;

4th categorical distance = 1/2 to 1 miles; and

5th categorical distance = 1 to 2 miles;

Note that the width of the outermost categorical distance is coincident with the value of D (*e.g.*, 2 miles). Figure 5 shows a sample computer generated map of Philadelphia with observational data appearing as point themes and with categorical

distance ranges represented as concentric circles. Figure 6 is a zoomed in view of the map of Philadelphia shown in Figure 5.

In step 220, beginning with the first observation (*i.e.*, neighboring property sale) in the data, the algorithm will identify the geographic location of the subject property by reading its X-Y coordinates. In step 230, moving to the next observation, the algorithm will compute the Euclidean distance to the first (*i.e.*, *k*th) observation. In step 240, if the distance is greater than *D*, then the program will return to step 230 via step 280 to retrieve the next observation. If, on the other hand, the distance computed in step 230 is less than or equal to *D*, the process will move to step 270.

In step 270, given the exact distance to the first subject property, the algorithm will flag the observation as lying within one of the particular *j* categorical distances. For example, if the distance from the first property to the second property is .87 miles, then the second property will be flagged as lying within the 4th categorical distance of 1/2 to 1 mile. This flag denotes that this property should be used in the computation of average property values that lie within 1/2 to 1 mile from the subject property.

In step 290, the algorithm then determines whether all (*i e* , *N*) observations in the dataset have been considered. If the last observation has not been considered, then the algorithm returns to step 230 via step 280 to evaluate the next observation. If, on the other hand, all observations have been considered such that every observation in the dataset has either been flagged as lying within one of the *j* categorical distances from the first subject property, or been ignored for being a distance greater than *D* from the first property, the algorithm moves to step 310.

In step 310, after having identified those properties that lie within particular

categorical distance from the subject property, the algorithm computes the average transaction price for each f the j categorical distances. In step 320, the algorithm populates the fields in the data with this value. For example, if five properties are identified as lying within 1/4 to 1/2 miles from the subject property, then the algorithm simply sums their transaction prices and divides by five. The algorithm then populates the field "Avg_1_2" with the calculated value.

In step 330, after having computed the average transaction price for each categorical distance, and populated the appropriate field with those values, the algorithm then ensures that it has considered every observation (*i.e.*, N) in the dataset. If every observation has not been considered, the algorithm returns to step 220 to retrieve the next dataset and declares it the subject property. If, on the other hand, every observation has been considered, the algorithm stops at step 360. The resulting product is a new text file that is identical to the original, except that several new fields have been added and populated with the values of average sales of surrounding properties.

Figure 4 shows the input and output data set (*i.e.*, text file) of the spatial algorithm process depicted (as detailed in Figure 3) given a three observation input file. As shown in Figure 4, the input data set is a three column table detailing, for example, latitude, longitude and sales price. The output data set is a eight column table (depending on the number of categorical distances, j) that details the average sales price of properties within each of the categorical distances.

In sum, the present invention provides a method for predicting a property transaction price for a prospective transaction location as a function of the attributes of the subject property and neighboring property sales. In particular, neighboring property sales

are evaluated using a spatial algorithm that computes the average price of the neighboring properties for different categorical distances. In this way, a regression coefficient may be applied to each categorical distance so as to account for the diminishing influence of neighboring properties as their distance from the subject property increases. The resultant equation may be incorporated into a computer system coupled to an observational database, so that the calculation is automated when a subject property location is entered.

It will be readily apparent that while the above description details one embodiment of the present invention, the present invention is by no means limited thereto. For example, from the above description, one of ordinary skill could implement the present invention on an Internet server coupled to an observational database. Clients connected to the Internet could submit a prospective transaction location, upon which the server would return a transaction price prediction. Of course, assuming national coverage, this implementation would require maintenance of a relatively large database.

What is claimed is:

1. A method for predicting a property transaction price for a prospective transacti n location, comprising:

determining a first value based on one or more attributes of said prospective transaction location;

determining a second value based on a relation of said prospective transaction location to one or more other properties; and

predicting said property transaction price for said prospective transaction location as a function of said first value and said second value.

2. The method of claim 1, wherein said step of determining said second value comprises:

determining one or more categorical ranges relative to said prospective transaction location;

computing average property values within each categorical range using a spatial average algorithm;

weighting each of said average property values based on its associated categorical range; and

adding said weighted average property values.

3. The method of claim 2, wherein said categorical ranges represent ranges of distance from said prospective transaction location.

4. The method of claim 3, wherein said categorical ranges represent discrete and sequential ranges of distance.

5. The method of claim 3, wherein said weighting is based on a distance of said categorical range from said prospective transaction location.

6. The method of claim 2, wherein said average property values are based on prior property sales.

7. The method of claim 1, further comprising determining which of said one or more properties are relevant to said prospective transaction location.

8. The method of claim 7, wherein said relevant property sales are based on a distance of said property sale from said prospective transaction location.

9. The method of claim 1, wherein said step of determining said first value comprises:

selecting said one or more attributes of said prospective transaction location;

determining an attribute value for each of said attributes of said prospective transaction location; and

adding said attribute values.

10. The method of claim 1, wherein said attributes of said prospective transaction location include square footage of a lot, size and style of said location, number of bedrooms, amenities of said location, characteristics of said location's neighborhood.

11. A computer-readable medium having computer-executable instructions for predicting a property transaction price for a prospective transaction location, comprising:

determining a first value based on one or more attributes of said prospective transaction location;

determining a second value based on a relation of said of said prospective transaction location to said one or more other properties; and

predicting said property transaction price for said prospective transaction location based on said first value and said second value.

12. The computer-readable medium of claim 11, wherein said step of determining said second value comprises:

storing one or more parameters of one or more property sales;

determining one or more categorical ranges relative to said prospective transaction location;

computing average property values within each categorical range using a spatial average algorithm;

weighting each of said average property values based on its associated categorical range; and

adding said weighted average property values.

13. The computer-readable medium of claim 12, wherein said categorical ranges represent ranges of distance from said prospective transaction location.

14. The computer-readable medium of claim 13, wherein said categorical ranges represent discrete and sequential ranges of distance.

15. The computer-readable medium of claim 13, wherein said weighting is based on a distance of said categorical range from said prospective transaction location.

16. The computer-readable medium of claim 11, wherein said average property values are based on prior property sales.

17. The computer-readable medium of claim 11, having further computer-executable instructions that determine which of said plurality of property sales are relevant to said prospective transaction location.

18. The computer-readable medium of claim 17, wherein said relevant property sales are based on a distance of said property sale from said prospective transaction location.

19. A computer-readable medium having computer-executable instructions for predicting a property transaction price for a prospective transaction location, comprising the sum of a first value and a second value, wherein said first value is based on an attribute

value for one or more attributes of said prospective transaction location, and wherein said second value is based on a sum of weighted average property values, each of said average property values corresponding to a categorical range relative to said prospective transaction location, and each of said average property values computed using a spatial average algorithm.

20. The computer-readable medium of claim 19, wherein said weighting of said average property values is based on said corresponding categorical range of said average property value.

21. The computer-readable medium of claim20, wherein said categorical ranges represent discrete ranges of distance from said prospective transaction location.

22. The computer-readable medium of claim19, wherein said one or more of said attributes of said prospective transaction location include square footage of a lot, size and style of said location, number of bedrooms, amenities of said location, characteristics of said location's neighborhood.

ABSTRACT

The present invention provides a method and computer-executable instructions on a computer-readable medium for predicting a property transaction price for a prospective transaction. The method determines a first value and a second value and predicts the property transaction as a function of those values. The first value is based on one or more attributes of the prospective transaction location and the second value is based on a relation of the prospective transaction location to one or more other properties.